

Figure 1

Nucleotide sequence of the *Prunus amygdalus HNL5* gene obtained by PCR amplification

The start codon (ATG) and stop codon of the open reading frame are printed in bold type, and the nucleotides in the intron regions are indicated in lower case letters. The peripheral sequences which have been attached via the PCR primers and which are not part of the *HNL5* gene are underlined. The splice sites of the introns were identified with the aid of the consensus sequence "GT....AG".

1 GGAATTCACA ATATGGAGAA ATCAACAA**ATG** TCAGTTATA**C**T TTTTG**TGTT**
51 GCATCTTCTT **GTTCTTCATC** TTCAGTATT**C** AGAGGTT**CAC** TCGCTTG**CCA**
101 ATACTTCTGC TCATGgtaaa **t**tcacat**t** cagtatt**c**at **t**taacagcaa
151 aatggtgaga **t**tataat**a**ta agaa**a**actga cacaagg**at**g gcaag**aa**aca
201 agctaatt**a**ta galgc**a**tgt**t** gaaaaaa**a**tc **t**tcacat**c**tc t**c**acatata**t**
251 **t**tgcagATT TTAGCTACTT GAAGTT**TG**T TACAA**CG**CCA CTGATAC**AAG**
301 CTCGGAAGGA TCATATGACT ACATTGTAAT CGGTGGAGGA ACATCAGGG**T**
351 GTCCATTGGC AGCAACT**TTA** TCAGAAAAAT ACAAGGT**GCT** TCTTCTAGAA
401 AGAGGC**ACTA** TTG**C**TACAGA ATACCCGAAC ACGTT**GACTG** CAGATGGG**TT**
451 TGCATATAAT CTGCAGCAAC AAGATGAT**GG** AAAGACGCCA GTT**GAA**AGGT
501 TCGTGTCCGA AGATGGCATT GATAATGTGC GAGCCAGGAT CCTCGGTGGC
551 ACGACCATAA TCAATGCAGG CGTCTACGCC AGAGCTAAC**A** TTT**CATT** CTA
601 TAGTCAAACA GGAATTGAAT GGGACCTGGA TTTGGTCAAT AAGACAT**ATG**

651 AGTGGGTTGA AGACGCCATT GTGGTCAAGC CAAATAATCA ATCTTGGCAA
701 TCTGTTATAG GAGAGGGATT CTTGGAGGCG GGTATTCTTC CAGACAATGG
751 ATTTAGTTTG GATCACGAAG CAGGAACTAG ACTCACCGGC TCAACTTTG
801 ACAATAATGG AACGCGACAT GCGGCTGATG AACTGCTTAA TAAAGGAGAC
851 CCTAATAACT TGCTAGTTGC AGTCAGGCC TCAGTAGAGA AGATCCTCTT
901 CTCTTCCAAT ACATCAAgtatgttgcatca gtgtatattttatgttgtatgc
951 ctatgttgtatgttgtatgcact cgaaaatttatattttatca tttttaaaata
1001 ctaacagaat agtgtgaagt ctcatatttc cttccatattttccaaatt
1051 tccataaaaca aaacttccca atttccttc gtttagttt gcaataatata
1101 taagctatttc tctaattgcag ATTTGTCAGC TATTGGAGTC ATATATACGG
1151 ATTCTGATGG AAACTCTCAT CAGGCATTG TACGCGGTAA CGGAGAAGTT
1201 ATTGTTAGTG CAGGGACAAT CGAACGCCT CAGCTTCTAC TACTTAGTGG
1251 CGTTGGACCA GAGTCTTACC TATCTTCTCT CAACATCACA GTTGTTCAGC
1301 CGAATCCTTA TGTTGGGCAG TTTGTGTATG ACAATCCTCG TAATTTCAATT
1351 AATATTTGC CCCCAAATCC AATTGAAGCC TCTGTTGTAA CTGTTTTAGG
1401 CATTAGAAAGT GATTATTATC AAGTTCTCT GTCAAGCTTGC CCATTTCCA
1451 CTCCACCCCTT TAGTCTTTT CCTACAAACAT CTTACCCCTT CCCAAATTG
1501 ACTTTGCTC ATATTGTTAG CCAAGTTCCA GGACCATTGT CTCATGGTTC
1551 TGTCACGCTA AATTCAATCAT CTGACGTGAG AATCGCTCCA AATATTAAAT
1601 TCAATTACTA TTCAAATTCC ACAGACCTTG CTAATTGTGT TAGCGGCATG
1651 AAGAAGCTTG GTGACTTATT AAGGACAAAG GCATTAGAAC CATATAAAGC

1701 TCGAGATGTG CTGGGAATTG ACGGTTCAA TTATTTGGGA GTACCTTGC
1751 CAGAGAACCA AACAGATGAT GCATCCTTCG AAACATTTG TCTAGATAAT
1801 GTAGCTTCAT ACTGGCATT CCACGGTGG AGCCTGTTG GGAAAGTGCT
1851 TGATGACAGT TTCCGTGTTA TGGGGATCAA AGCATTACGC GTTGTGATG
1901 CCTCCACTTT CCCTTACGAA CCAAACAGCC ATCCTCAGGG CTTCTATCTG
1951 ATGTTAGGAA Ggtatgtat gcacacttcc aaccactaga gatctcaat
2001 attttgttgtl tggtaatg aactcttcgc cgcatcgctc tttttattta
2051 atccctaaaaa ttttgttt tgccagGTA TGTGGGCCTT CAAATCCTGC
2101 AAGAAAGGTC AATCCGGTTG GAGGCTATTCA ATAATATTCA AGAGTCCATG
2151 TGAAGAATT CG

Figure 2

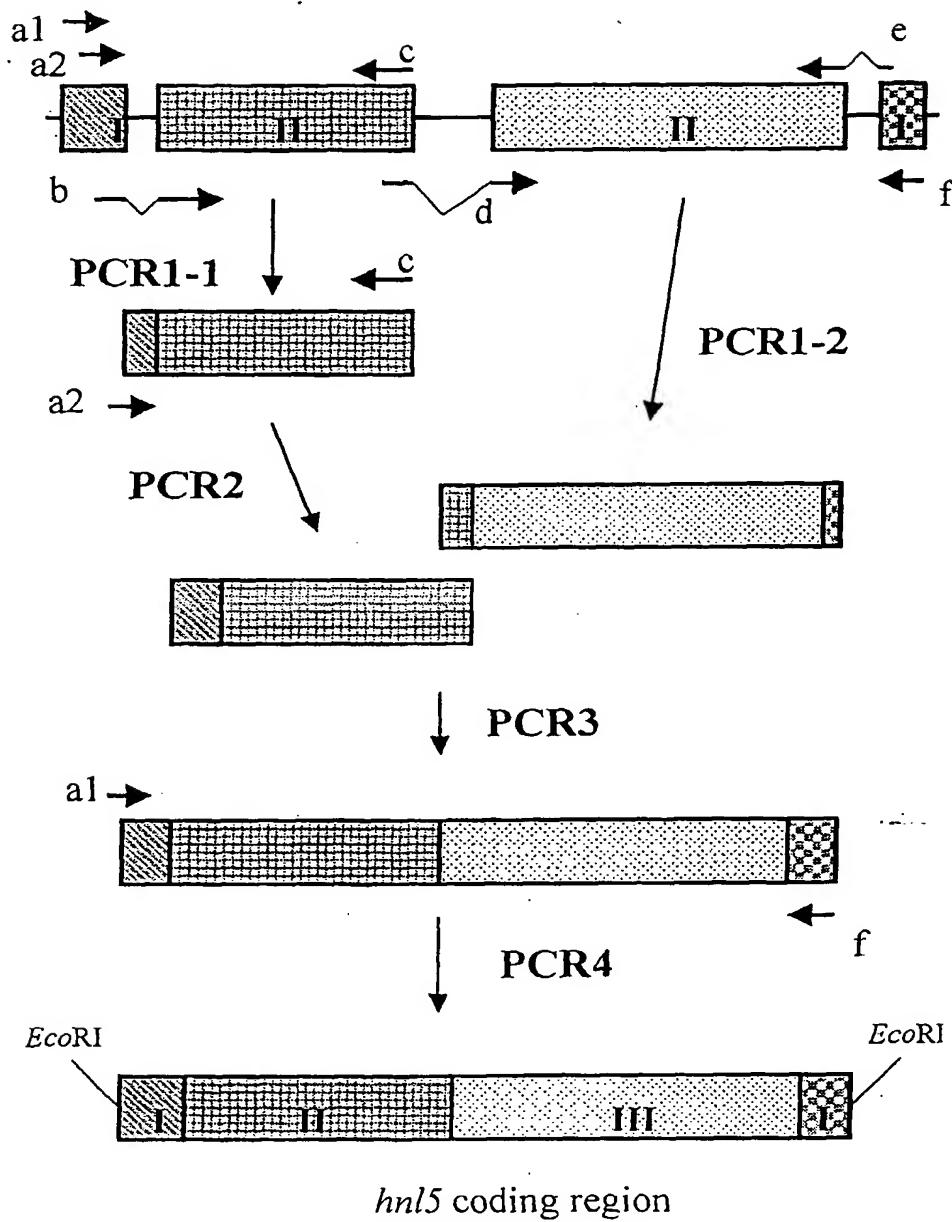


Figure 3:

Amino acid sequence of the *Prunus amygdalus* hydroxynitrile lyase (HNL5), derived from the nucleotide sequence of the *HNL5* gene. The signal sequence determined from sequence analysis is printed in bold type and the postulated processing site is indicated by an arrow. Possible glycosylation sites (PROSITE patterns) are underlined.



MEKSTMSVILFVLHLLVLHLQYSEVHSLANTSAHDFSYLKFVYNATDSSEGSDYI****
VIGGGTSGCPLAATLSEKYKVLLERGTIAEPNTLTADGFAYNLQQQDDGKTPVE
RFVSEDGIDNVRARILGGTTIINAGVYARANISFSQTGIEWDLDLNKTEWVEDAI
VVKPNNQSWQSVIGEGFLEAGILPDNGFSLDHEAGTRLTGSTFDNNGTRHADELL
NKGDPNLLVAVQASVEKILFSSNTSNLSAIGVITDSDGNSHQAFVRGNEVIVSA
GTIGTPQLLLSGVGPESYLSSLNITVQPNPYVGQFVYDNPRNFINILPPNIEASVV
TVLGIRSDYYQVSLSSLPFSTPPFSLFPTTSYPLPNSTFAHIVSQVPGPLSHGSVTLN
SSSDVRIAPNIKFNYYSSNSTDLANCVSGMKKLGDLLRTKALEPYKARDVLGIDGFNY
LGVPLPENQTDDASFETFCLDNVASYWHYHGGSLVGKVLDDSFRVMGIKALRVVD
ASTFPYEPNSHPQGFYMLGRYVGLQILQERSIRLEAIHNIQESM

20046232 041602

Figure 4:

Nucleic acid sequence of the DNA fragment coding for a secretory hybrid protein (PamHNL5xGOX) with HNL activity, consisting of sequences of the *Prunus amygdalus* HNL5 gene and the *Aspergillus niger* glucose oxidase gene.

gaattcatcatgcagactctccgtggctcgctgtggcgtccccgtgcggccctgcca-
caactacatcaggagcaatggcattgaaggctacaacgccactgataacaagctcggaaaggatca-
tatgactacatgtaatcggtgaggaaacatcagggtgtccattggcagcaactttatcagaa-
aaataacaagggtgcitctttagaaagaggcactattgtacagaatacccaaacacgtt-
gactgcagatgggttgcataatctgcagcaacaagatgtggaaagacgccagttga-
aagggtcggtccgaagatggcattgataatgtgcgagccaggatccctggcggcacgacca-
taatcaatgcaggcgtctacgccagagctaacaatttcattctatagtcaaacaggaatt-
gaatgggacctggatttgtcaataaagacataatgagtgggtgaagacgccatgtgg-
caagccaaataatcaatctggcaatctgttataggagaggattctggaggcggg-
tattcttccagacaatggatttagtttgatcacgaagcaggaactagactcaccggct-
caacatttgacaataatggAACGCGACATGCCGTGATGAACTGCTTAATAAAGGAGACCC-
taataacttgclagttgcagttcaggcctcagtagagaagatcctctcttccaatacat-
caaatttgtcagctattggagtcatatatacggattctgtatggaaactctcatcaggcatttg-
tacgcggtaacggagaagttttgttagtgcagggacaatcggAACGCCCTCAGCTTAC-
tacttagtggcgttgtggaccagagtcacatctttcaacatcacagtgttcagcc-
gaatcccttatgtgggcagtttgtgtalacaatccctgtatttcatatattttggccc-
caaattccaattgaaggccttgtgtactgttttagcattagaagtgtattat-
caagtttctgtcaagcttgccatttccactccaccctttagtctttccataacatct-
tacccctcccaaattcgactttgtcatattgttagccaagttccaggaccattgtct-
catgggtctgtcacgtcaaattcatcatctgacgtgagaatcgtccaaatattaaattcaat-
tactattcaaaattccacagacccatgtcaatttgttagcggcatgaagaagcttgtgact-
tattaaggacaaaaggcattagaaccataaaagctcgagatgtgtggaaatgtacggttt-
caattttgggagttacccttgccagagaaccaaacagatgtatgcaccccttcgaaa-
cattttgtctagataatgttagctcatactggcattaccacccgtggaaaggcccttgtggg-
aagtgcgtgtacagttccgttgtatgggatcaaaggcattacgcgttgtgtatgcctc-
cactttcccttacgaaccaaacagccatccctaggcattctatgtatgttaggaagg-...
tatgtgggccttcaaattcctgcaagaaaggtaatgcagttgcggccgtgcgaatttc

Figure 5: Amino acid sequence of the hybrid protein PamHNL5xGOX, derived from the nucleic acid sequence (figure 4).

MQTLLVSSLVSLAAALPHYIRSNGIEAYNATDSSEGSYDYIVIGGGTSGCPLAATL
SEKYKVLLERGTIATEYPNTLTADGFAYNLQQQDDGKTPVERFVSEDGIDNVRARI
LGGTTIINAGVYARANISFYSQTGIEWDLVLNKTYEWVEDAIVKPNNQSWQSVIG
EGFLEAGILPDNGFSLDHEAGTRLTGSTFDNNGTRHADELLNKDPNNLLVAVQA
SVEKILFSSNTSNLSAIGVIYTDSDGNSHQAFVRGNGEVIVSAGTIGTPQLLLSGVG
PESYLSSLNITVQPNPYVGQFVYDNPRNFINILPPNPIEASVTVLGIRSDYYQVSLS
SLPFSTPPFSLFPTTSYPLPNSTFAHIVSQVPGPLSHGSVTLNSSSDVRIAPNIKFNY
YSNSTDLANCVSGMKKLGDLLRTKALEPYKARDVLGIDGFNYLGVPLPENQTDDAS
FETFCLDNVASYWHYHGGSLVGKVLD DSFRVMGIKALRVVDASTFPYEPNSHPQG
FYLMMLGRYVGLQILQERSMO

200044222 • 044602

Figure 6: Comparison of the amino acid sequences of *Prunus amygdalus* HNL5 and of the hybrid protein PamHNL5xGOX. Sequence parts of *Aspergillus niger* glucose oxidase are underlined. Sequence regions having no significant homology between the two proteins are printed in italics, and the signal peptides are printed in bold type.

PamHNL5Gox 1 mgtlvsslv vslaa^{alpha}lphy lsngiea-- ----- --YNATDTSS

PamHNL5 1 mekstmsvil fvlhllv*lh* qysevh*s*lan tsahd*s*ylk fvYNATDTSS

PamHNL5Gox 37 EGSYDYIVIG GGTSGCPLAA TLSEKYKVLL LERGTIALEY PNTLTADGFA

PamHNL5 51 EGSYDYIVIG GGTSGCPLAA TLSEKYKVLL LERGTIALEY PNTLTADGFA

PamHNL5Gox 87 YNLQQQDDGK TPVERFVSED GIDNVRARIL GGTTIINAGV YARANISFYS

PamHNL5 101 YNLQQQDDGK TPVERFVSED GIDNVRARIL GGTTIINAGV YARANISFYS

PamHNL5Gox 137 QTGIEWDLDL VNKT^YEWWED AIVVKPNNQS WQSVIGEGFL EAGILPDNGF

PamHNL5 151 QTGIEWDLDL VNKT^YEWWED AIVVKPNNQS WQSVIGEGFL EAGILPDNGF

PamHNL5Gox 187 SLDHEAGTRL TGSTFDNNGT RHADELLNK GDPNNLLVAV QASVEKILFS

PamHNL5 201 SLDHEAGTRL TGSTFDNNGT RHADELLNK GDPNNLLVAV QASVEKILFS

PamHNL5Gox 237 SNTSNLSAIG VIYTDSDGNS HQAFVRGNGE VIVSAGTIGT PQLLLLSGVG

PamHNL5 251 SNTSNLSAIG VIYTDSDGNS HQAFVRGNGE VIVSAGTIGT PQLLLLSGVG

PamHNL5Gox 287 PESYLSSLNI TVVQPNPYVG QFVYDNPRNF INILPPNPIE ASVVTVLGIR

PamHNL5 301 PESYLSSLNI TVVQPNPYVG QFVYDNPRNF INILPPNPIE ASVVTVLGIR

PamHNL5Gox 337 SDYYQVSLSS LPFSTPPFSL FPTTSYPLPN STFAHIVSQV PGPLSHGSVT

PamHNL5 351 SDYYQVSLSS LPFSTPPFSL FPTTSYPLPN STFAHIVSQV PGPLSHGSVT

PamHNL5Gox 387 LNSSSDVRIA PN^IKFNYYSN STDLANCVSG MKKLGDLLRT KALEPYKARD

PamHNL5 401 LNSSSDVRIA PN^IKFNYYSN STDLANCVSG MKKLGDLLRT KALEPYKARD

PamHNL5Gox 437 VLGIDGFNYL GVPLPENQTD DASFETFCLD NVASYWHYHG GSLVGKVLD

PamHNL5 451 VLGIDGFNYL GVPLPENQTD DASFETFCLD NVASYWHYHG GSLVGKVLD

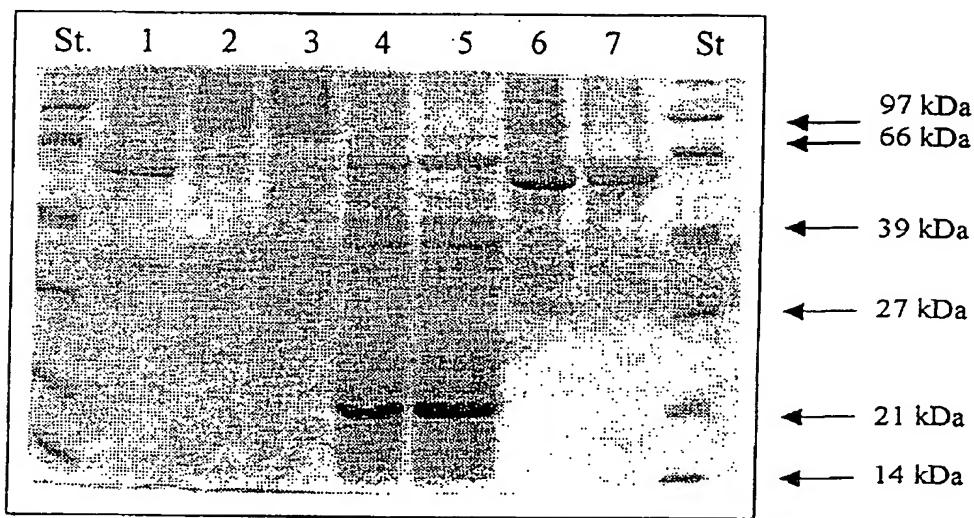
PamHNL5Gox 487 SFRVMGIKAL RVVDASTFPY EPNSHPQGFY LMLGRYVGLQ ILQERS*mg-*

PamHNL5 501 SFRVMGIKAL RVVDASTFPY EPNSHPQGFY LMLGRYVGLQ ILQERS*irle*

PamHNL5Gox 535 -----

PamHNL5 551 *aihniqesm*

Figure 7: Analysis of HNL preparations by SDS PAGE.
Details are described in example 11.



10045232 011602

Figure 8:

Nucleotide sequence of the *Prunus amygdalus HNL1* gene obtained by PCR amplification.

ATGGAGAAATCAACAATGTCAGCTATACTGTTGGTGTATACATTTGTCCTCC
ATCTTCAATATTCTGAGGTCCACT
CGCTTGCCACGACTCTGATCATGgtaaatcactcaaccgttaatcaaaaacacaaaaagg-
caatcaaaaagaaaaacg
gaaaaaaagtgtaaagaaaaagcagatatacgccctgcatacatgtgtctatatactttaaaaactcttcgtctt
gagatttgtagATTTAGCTACCTGAGCTTGCATACGACGCCACTGATCTA-
GAGTTGGAAGGATCATATGACTACGT
TATAGTTGGCGGAGGAACATCAGGGTGTCCATTGGCAGCAACTTATCAGAAAA
ATACAAGGTGCTCGTTCTCGAAAGG
GGCAGTCTCCGACAGCATATCCCAACGTCTGACTGCAGATGGTTGTATAT
AATCTCCAGCAAGAAGATGATGGAA
AGACACCAGGTCGAAAGGTTGTCGAAAGATGGTATTGATAATGTACGGGC
AGGGTGCTCGGTGGCACAAGCATTAT
CAATGCCGGTGTCTACGCCAGAGCTAACACCTCAATCTAGTGCATCAGGAGT
TGATTGGACATGGATTGGTTAAT
CAGACATATGAGTGGTTGAAGACACTATTGTGTACAAGCCAATTCTCAATCTT
GGCAGTCTGTTACAAAAACTGCAT
TCTTGAGGCTGGTGTTCATCCAAACCATGGATTAGTTAGATCATGAAGAAG
GAACAGAATTACCGGCTCAACTTT
TGACAACAAGGGAACGAGACATGCAGCTGATGAACCTTAATAAGGAAACTC
TAACAACCTGCGAGTTGGAGTTCAT
GCCTCAGTAGAGAAGATCATCTTCTCCAATGCACCAAGgtatgtgcatacgactccaa-
aattaatatttgtcaatt
taaaaactagcaggagccaaggctggaaagtacgaataaaattttttccctggatttgtataatgatla
taagctttctgtatgttagGTTTGACAGCTACAGGAGTCATATAGGGATTCTAATG-
GAACGCCTCACCAAGCATT
GTACGCAGTAAGGGAGAAGTTATCGTGAGTGCAGGGACAATTGGGACCCCTCA
ACTTCTACTACTTAGCGGTGTTGGGC

2004-09-23 04:16:02

CAGAGTCTTACCTATCATCTCTAAATATTCCAGTTGTTCTTCCCATCCTACGTC
GGACAGTTCTGCATGACAATCC
TCGTAATTCACTAACATTGGCCCCAAATCCAATTGAACCCACAATTGTAACTG
YTCTAGGCATTTCAAACGATTTC
TACCAATGTTCTTCTCGAGCTGCCATTACAACCTCCACCCCTCGGTTTTCC
CTAGTGCATCTTATCCCCTGCCAA
ATTCGACTTTGCTCACTTGCTAGCAAAGTGGCAGGACCTTATCATATGGTTC
TCTCACACTGAAATCATCCTCCAA
TGTGAGAGTCAGTCAAATGTCAAATTAAATTACTATTCAAATCTGACAGATCTT
CTCATTGTGTTAGCGGCATGAAG
AAGATTGGTGAACCTTGAGCACAGACGCATTAAAACCATAAAAGTTGAAGATT
TGCCGGGTGTTAGAAGGTTTAATA
TTTGGGAATCCCTTGCCAAAGGACCAAACAGATGATGCAGCCTCGAAACAT
TTGCCGAGAATCAGTAGCCTCATA
TTGGCACTACCACGGTGGATGCCTTGGAAAGGTGCTTGATGGTGATTCCG
TGTACAGGGATCAATGCATTACGC
GTTGTTGATGGCTAACATTCCCTACACACCAGCGAGCCACCCCTCAGGGCTTC
TATCTGATGTTAGGGAGgtatgta
caaattctcaataattttgggtgagtggcttgttgcataatgaactctatgccatatttcatttcattccatccattcca
ttttgtgccatggcagGTATGTGGCATTAAAATTCTGCAAGAAAGATCAGCTTCA-
GATCTAAAATCTGGATTCC
CTCAAGTCAGCAGCATCCTGGTTCTTAAACT

1004632-041602

Figure 9:

Amino acid sequence of *Prunus amygdalus* hydroxynitrile lyase (HNL1), derived from the nucleotide sequence of the HNL1 gene.

MEKSTMSAILVLYIFVLHLQYSEVHSLATTSDHDFSYLSFAYDATDLELEGSY
DYVIVGGGTSGCPLAATLSEKYKVLVLERGSIPTAYPNVLADGFVYNLQQE
DDGKTPVERFVSEDGIDNVRGRVLGGTSIINAGVYARANTSISASGVWDWM
DLVNQTYEWVEDTIVYKPNSQSWSVTKTAFLLEAGVHPNHGFSLDHEEGTRI
TGSTFDNKGTRHADELLNKGNSSNLRVGVHASVEKIIFSAPGLTATGVIYR
DSNGTPHQAFVRSKGEVIVSAGTIGTPQLLJSGVGPESYLSNLIPVVLSPY
VGQFLHDNPRNFINILPPNPIEPTIVTLGISNDYQCSFSSLPPFTPPFCFFPS
ASYPLPNSTFAHFASKVAGPLSYGSLTLSKSSNVRSVPNVKFNYYSNLDSHC
VSGMKKIGELLSTDALKPYKVEDLPGVEGFNILGIPLPKDQTDDAAFETFCR
ESVASYWHYHGGCLVGKVLDGDFRVTGINALRVVDGSTFPYTPASHPQGFYL
MLGRYVGIKILQERSASDLKILDSDLKSAASLV

40046232-014602